Computers and Electronics in Agriculture 205 (2023) 107585

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

In-field classification of the asymptomatic biotrophic phase of potato late blight based on deep learning and proximal hyperspectral imaging

Chao Qi [a,b], Murilo Sandroni [c], Jesper Cairo Westergaard [d], Ea Høegh Riis Sundmark [e], Merethe Bagge [e], Erik Alexandersson [c], Junfeng Gao [a,f,*]

[a] Lincoln Agri-Robotics, Lincoln Institute for Agri-Food Technology, University of Lincoln, Lincoln, UK
[b] College of Engineering, Nanjing Agricultural University, Nanjing 210031, China
[c] Department of Plant Protection Biology, Swedish University of Agricultural Sciences, Alnarp, Sweden
[d] Department of Plant and Environmental Sciences, University of Copenhagen, Taastrup, Denmark
[e] Danespo Breeding Company, Give, Denmark
[f] Lincoln Centre for Autonomous Systems, University of Lincoln, Lincoln, UK

ARTICLE INFO

Keywords:
Asymptomatic biotrophic phase
Wavelength selection
Attention networks
Convolutional neural networks
Plant phenotyping

ABSTRACT

Effective detection of potato late blight (PLB) is an essential aspect of potato cultivation. However, it is a challenge to detect late blight in asymptomatic biotrophic phase in fields with conventional imaging approaches because of the lack of visual symptoms in the canopy. Hyperspectral imaging can capture spectral signals from a wide range of wavelengths also outside the visual wavelengths. Here, we propose a deep learning classification architecture for hyperspectral images by combining 2D convolutional neural network (2D-CNN) and 3D-CNN with deep cooperative attention networks (PLB-2D-3D-A). First, 2D-CNN and 3D-CNN are used to extract rich spectral space features, and then the attention mechanism AttentionBlock and SE-ResNet are used to emphasize the salient features in the feature maps and increase the generalization ability of the model. The dataset is built with 15,360 images (64x64x204), cropped from 240 raw images captured in an experimental field with over 20 potato genotypes. The accuracy in the test dataset of 2000 images reached 0.739 in the full band and 0.790 in the specific bands (492 nm, 519 nm, 560 nm, 592 nm, 717 nm and 765 nm). This study shows an encouraging result for classification of the asymptomatic biotrophic phase of PLB disease with deep learning and proximal hyperspectral imaging.

1. Introduction

Disease monitoring of potatoes is an essential step to improve breeding and production. Late blight caused by the oomycete *Phytophthora infestans* is a devastating disease in potato production, affecting farmers' income and having a negative impact on the environment relying sometimes on weekly spraying of fungicides (Chen et al., 2021; Zheng et al., 2021), which are expensive and time-consuming to apply (Wang et al., 2021). Hence, accurate classification of the asymptomatic biotrophic phase is critical to screen for breeding material for resistance, reduce fungicide applications and effectively control PLB.

The development of low-cost sensor technologies for computer vision and remote sensing is paving the way for image-based agricultural management and show great potential for automatic detection of crop disease and diagnosis based on different types of images (e.g., RGB images (Su et al., 2021), thermal images (Yang et al., 2021), remote sensing images (Bagheri, 2020). Our previous work (Gao et al., 2021) has shown great potential to use visual images with deep learning for the evaluation of in-field PLB infestation based on the number of recognized lesions. However, it only focuses on advanced stages of PLB development and is unable to recognize pre-symptomatic potato plants infested with *P. infestans*. Hyperspectral imaging (HSI) benefits from capturing many and narrower spectral bands in a continuous spectral range. It provides two-dimensional spatial information and rich spectral information in the third dimension.

With the rapid development of hyperspectral remote sensing, many excellent machine learning-based algorithms have been proposed to solve hyperspectral image classification problems (Han and Deng, 2018 Zhang et al., 2019), such as Support Vector Machine (SVM), polynomial logistic regression, sparse representation, and cooperative representation. Rodriguez et al. (2021) proposed an Unmanned aerial vehicle (UAV)-based multispectral image detection method for PLB, evaluating

* Corresponding author at: Lincoln Agri-Robotics, Lincoln Institute for Agri-Food Technology, University of Lincoln, Lincoln, UK.

https://doi.org/10.1016/j.compag.2022.107585
Received 1 January 2022; Received in revised form 6 October 2022; Accepted 18 December 2022
Available online 31 December 2022

the performance of five machine learning algorithms: random forest, gradient augmented classifier, support vector classifier, linear support vector classifier, and k-nearest neighbor classifier to detect PLB with an accuracy of 0.982, 0.895, 0.979, 0.981, respectively. In controlled conditions, Gold et al. (2020) used random forest discrimination (RF), partial least squares discriminant analysis (PLS-DA), and normalized difference spectral index to detect regions of late blight occurrence with an accuracy of 70.94 %, 71.13 %, and 72.23 %, respectively. Traditional machine learning methods have some limitations, and the relatively simple mapping structure leads these methods to extract only shallow semantic image features.

To further explore the potential of deep learning in hyperspectral image classification, some algorithms with more advantageous generalization performance have been proposed. Representative models based on deep learning such as convolutional neural networks (CNNs), deep confidence networks, recurrent neural networks, and graph convolutional networks are suggested (Ma et al., 2020; Meng et al., 2021). CNN-based models can improve the classification of hyperspectral images of potato leaves because of their excellent performance in many image vision domains (Paoletti et al., 2018). Duarte-Carvajalino et al. (2018) proposed a custom 2D-CNN-based model designed to predict the severity of late blight impact in potato crops using multispectral remote sensing images and state-of-the-art machine learning algorithms, which achieved an accuracy of 0.74. Shi et al. (2021) proposed a new 3D-CNN deep learning model (CropdocNet) for accurate and automated late blight diagnosis with an accuracy of 94.2 %.

Although deep learning-based models can significantly improve the classification performance of hyperspectral images, most existing classification models have inherent restrictions, since they fail to extract sufficient spectral-spatial correlation information. The 2D-CNN-based deep network structure only utilizes spatial information, missing the important information from spectral signals. Similarly, the 3D-CNN-based network model performs poorly in many spectral bands to classify similar texture classes. The purpose of the method presented here is to detect the asymptomatic severity of PLB disease. During the biotrophic life stage of *P. infestans,* the degree of leaf infection cannot be judged by human visual inspection, thus the capability of extracting spectral-spatial features at this early stage of infection is critical. To tackle this, we propose a deep learning model combining 2D-CNN and 3D-CNN to make full use of the respective advantages of the two through customized feature extraction for the PLB disease detection. The main contributions of this paper are as follows:

1. A customized deep learning structure designed for classification of the asymptomatic biotrophic phase of PLB disease by combining 2D-CNN and 3D-CNN as well as attention networks.
2. The extraction of important bands for PLB classification and validating their classification effectiveness compared to classification using all wavelengths.
3. The verification of the superiority of model PLB-2D-3D-A by comparing with traditional machine learning methods and deep learning methods (2D-CNN and 3D-CNN).

The rest of the paper is organized as follows: Section 2 describes the dataset and the proposed model, Section 3 presents the experimental results, Section 4 provides a discussion and highlights future work, and Section 5 summarizes this work.

2. Materials and methods

2.1. Dataset

The data were collected on July 9, July 13, July 15, and July 18, 2020 as part of the Danespo field trial outside of Give, Denmark (N 55.800586, E 9.223748) and contained 20 different potato genotypes varying in susceptibility to PLB (four cultivars and sixteen breeding lines). Hyperspectral images with a resolution of 512*512*204 and a band of 204 were collected using a handheld hyperspectral camera, Specim IQ, from Specim (Oulu, Finland). Potatoplants from the infection rows were inoculated with *P. infestans* on 7th July 2020, from where the pathogen was expected to spread over the studied plots. *P. infestans* isolated from a local field the previous growth season were maintained at 4 °C. Before field inoculation propagation were done on disinfected tuber discs at 16 °C at 100 % humidity for 10 days. The development of mycelia and sporangia were examined in a light microscope. Ca 4 potato discs were then washed in 20L of 4 °C tap water and cold-treated at 4 °C for ca 1 h to release sporangia. In the field potato inoculation rows were then sprayed twice using a handheld pump to initiate the infection. Images collected on the third, seventh, ninth and twelfth day after inoculation were classified as separate categories. Among them, the images collected in the third, seventh and ninth days could not be discerned by the naked eye as there were no visual disease lesions, and we defined them as very early stage of late blight. While few of the images in the twelfth day could be discerned by the naked eye as disease lesions, we still consider them as early PLB. To reduce the computational burden, each collected raw image (512*512*204) were cropped into 64 subimages (64*64*204) to scale up training samples and reduce computation burden. To ensure the integrity of the spatial information, we did not pre-process the channel information. The number of images per genotype is 12 (We followed the same leaves throughout the infection), thus the total number of data set is 15,360 (64*12*20) and the number of each category is 3840. The examples of original images in each category are shown in Fig. 1.

2.2. Radiation calibrations

Radiation calibration (Behmann et al., 2018) mainly includes dark current removal from the original data radiation intensity, gain calibration of each spectral data channel and wavelength revision, etc. Assuming that the calibration is linear, we have

$$L = a^{*}DC + b \tag{1}$$

$$D = a^{*}DC_0 + b \tag{2}$$

where L refers to the radiance of the image at the sensor pupil ($W/(m^2 \cdot sr \cdot um)$) calculated using the radiative transfer model. D indicates the irradiance of the corresponding image in the dark current case ($W/(m^2 \cdot sr \cdot um)$). DC represents the grey scale value of the image element on the corresponding remote sensing image. DC_0 describes the dark current value of the image element on the corresponding remote sensing image. a and b are the calibration coefficients to be found. In the dark current case, D should be zero, and the calibration coefficients a and b can be found by combining Equations (3) and (4). Example of images calibrated by radiation are shown in Fig. 2.

$$a = L/(DC - DC_0) \tag{3}$$

$$b = -a \cdot DC_0 \tag{4}$$

2.3. Background removal

The purpose of background removal is to remove information from hyperspectral images of objects besides potato leaves. The result of the background removal is shown in Fig. 3.

$$\begin{cases} \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2} \geqslant \frac{1}{N}\sum_{i=1}^{N}x_i, & NaN \\ \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2} < \frac{1}{N}\sum_{i=1}^{N}x_i, & x_i \end{cases} \tag{5}$$

Where N represents the total number of pixels, x_i stands for the pixel

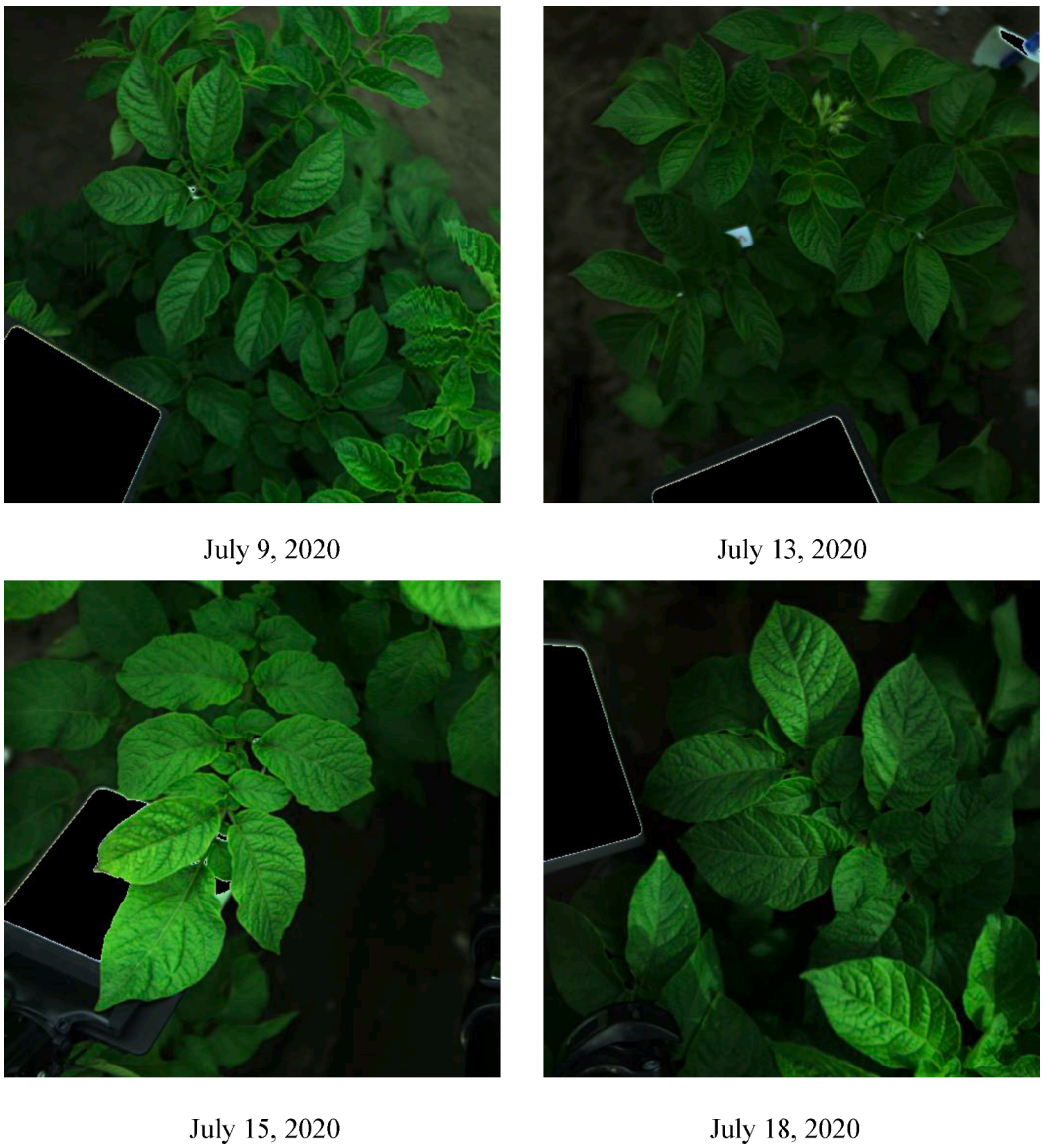


Fig. 1. Examples of the infected potatoes at four different dates, displaying with RGB bands.

value and μ indicates the pixel average.

2.4. Band screening

Although hyperspectral images provide a great deal of finer feature information, not all the images from each band could provide useful information. Redundancy in hyperspectral imagery can be both spatial and spectral. Spatial redundancy arises at the same band, as the greyscales of sampled points on the same feature surface are typically spatially coherent with each other, and feature greyscales based on discrete pixel sampling do not take full advantage of this feature. Spectral redundancy arises because the high spectral resolution and high data dimensionality of hyperspectral images allow information in one band of the image to be partially or fully predicted by other bands. More specifically, our dataset captures 204 bands, leading to that the parameter number to be processed reaches about 50 million for each image. Any redundant data brings a heavy computational burden, prone to slow or even failed network training. Therefore, it is necessary to verify whether screening out specific bands has an impact on the detection results. We use spectral first-order differentiation and spectral second-order differentiation to conduct band screening separately, and the final results are intersected to improve the quality of the band screening. The principles of spectral first order differentiation and spectral second order differentiation are as follows:

The derivative of a function is defined as:

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \tag{6}$$

While the h is small enough, we can use a centered difference formula to approximate the derivative:

$$f'(x_i) \approx \frac{f(x_i+h)-f(x_i-h)}{2h} \tag{7}$$

In practice, Origin treats discrete data by the transform of the centered difference formula, and calculates the derivative at point P_i by taking the average of the slopes between the point and its two closest neighbors.

The derivative function applied to discrete data points can therefore be written:

$$f'(x_i) = \frac{1}{2}\left(\frac{y_{i+1}-y_i}{x_{i+1}-x_i}+\frac{y_i-y_{i-1}}{x_i-x_{i-1}}\right) \tag{8}$$

When smooth option is chosen in differentiate, and X data is evenly spaced, Savitzky-Golay (1964) method will be used to calculate the

Fig. 2. Some images after radiometric calibration (For visualization, the images displayed here are composited with three RGB bands from hyperspectral images). The first row represents the images before the radiation calibration, and the second row shows the images after the radiation calibration.

Fig. 3. Results of background removal. The objects creating noise in the images include calibration plates, marker plates, tripods, etc. Since the images were collected in varying illumination, it is difficult to remove this background objects by selecting a specific threshold. Through experience pixels were removed when the standard deviation was greater than one half of the pixel mean. The equation for background noise removal is given in Equation (5).

derivatives.

First perform a polynomial regression on the data points in the moving window. The polynomial value at position × can be calculated as:

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0 \tag{9}$$

where n is the polynomial order, and $a_i, i = 0 \cdots n$ are fitted coefficients.

And first order derivative at position × is:

$$f'(x) = n a_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1 \tag{10}$$

Second order derivative at position × is:

$$f''(x) = n(n-1)a_n x^{n-2} + (n-1)(n-2)a_{n-1}x^{n-3} + \cdots + a_2 \tag{11}$$

2.5. Classification model

2.5.1. Plb-2D-3D-A

First, the pre-processed image (512*512*204) is uniformly sliced into 64*64*204 hyperspectral images using 3D convolution. To avoid the loss of spatial-spectral information, we only slice for pixel information instead of channel information. Second, to better investigate spatial-spectral features, this paper designed a sub-structure containing 2D-CNN and 3D-CNN. In this sub-structure, spatial feature information is extracted by 2D-CNN and spectral-spatial context information is obtained by 3D-CNN, and Batch Normalization (BN) (Ioffe and Szegedy, 2015) and Rectified Linear Activation Function (RELU) (Krizhevsky et al., 2012) are embedded after each convolution operation. Moreover, the extracted features are combined with two attention mechanisms, AttentionBlock and SE-ResNet, to fully extract spectral-spatial features and emphasize saliency features. Finally, the identifiable spatial-spectral features are fed into a 1 × 1 convolutional layer to facilitate

classification. The PLB-2D-3D-A structure is shown in Fig. 4. The experiments were conducted on a server with NVIDIA Tesla V100, CUDA 11.2, and Ubuntu 16. During the training process, the key hyperparameters were set as follows: dropout = 0.4, epoch = 60, batch size = 16, learning rate = $2e^{-4}$, and the optimizer was Adam (Kingma and Ba, 2014).

2.5.2. Spatial feature extraction

To extract feature maps with sufficient spatial-spectral contextual information, two-dimensional and three-dimensional convolution operations were designed in the PLB-2D-3D-A model. The following describes how to combine the 2D CNN and 3D CNN so that the feature maps comprise rich spatial-spectral correlation information. Before passing into the deep network, a neighborhood block $P \in \mathbb{R}^{S\times S\times B}$ is created by selecting a neighborhood square of size $S \times S$ around the center pixel from I, with the spatial location of the center at (α, β) and the number of bands B. In the experiments, S is set to 11 and B to 10.

For spatial correlation feature extraction, 2D-CNN operations are used as the basic unit for spatial feature extraction. The spatial extraction region contains four 2D-CNN $\left(con2D\_\{i\}_{i=1}^{4}\right)$ layers to capture feature maps of various spatial dimensions. After the conv2D_1, conv2D_2 and conv2D_4 operations, 3 × 3 convolution kernels with (1,1) downsampling steps are applied. conv2D_3 uses 2 × 2 convolution kernels with (2,2) steps.

For spectral relevant feature extraction, 3D convolution operations are employed to capture the correlation of the spectral dimensions. As shown in Fig. 4, four 3D convolutional layers $\left(conv3D\_\{i\}_{i=1}^{4}\right)$ operations are deployed to acquire spectral features of various depths. In conv3D_1, conv3D_2 and conv3D_4, the three 3D convolution kernels are 3 × 3 × 4, 3 × 3 × 4 and 3 × 3 × 2, respectively, where each convolution kernel has a downsampling step of (1,1,1). conv3D_3 contains convolution kernels of 2 × 2 × 2 and a step of (2,2,2).

The first three spatially relevant feature maps are fused with the spatial-spectral correlated feature maps so that the fused feature maps comprise rich relevant information. The fourth spatially correlated feature map is fused with the spectral-spatially correlated feature map to assist classification.

2.5.3. AttentionBlock

Based on the later experimental results, it can be observed that the PLB-2D-3D-A model including the attention module AttentionBlock (Yin et al., 2020) gives superior classification results. Further investigation identified the following steps explains the effectiveness of AttentionBlock from both theoretical and experimental perspectives. In the proposed model, AttentionBlock is utilized to emphasize the relevant information in the fused spectral space feature map. First, all the pixels in the feature map can be represented as $X_{in} = \{x_1, x_2, \cdots, x_N\}$, where N denotes the number of pixels. Each pixel x_i is represented as an E dimensional vector, where E refers to the channel number of the feature map. The output of this module is $O_{\text{out}} = \{o_1, o_2, \cdots, o_N\}$. The similarity between each two pixels x_i and x_j can be expressed as:

$$o_i = softmax\left(\phi(x_i)^T\varphi(x_j)\right)g(x_j) = \frac{1}{\Sigma_{\forall j} exp^{exp\left(\phi(x_i)^T\varphi(x_j)\right)}}\left(\phi(x_i)^T\varphi(x_j)\right)g(x_j) \tag{12}$$

Where $\phi(\bullet)$, $\varphi(\bullet)$ and $g(\bullet)$ all represent a Conv2D with a

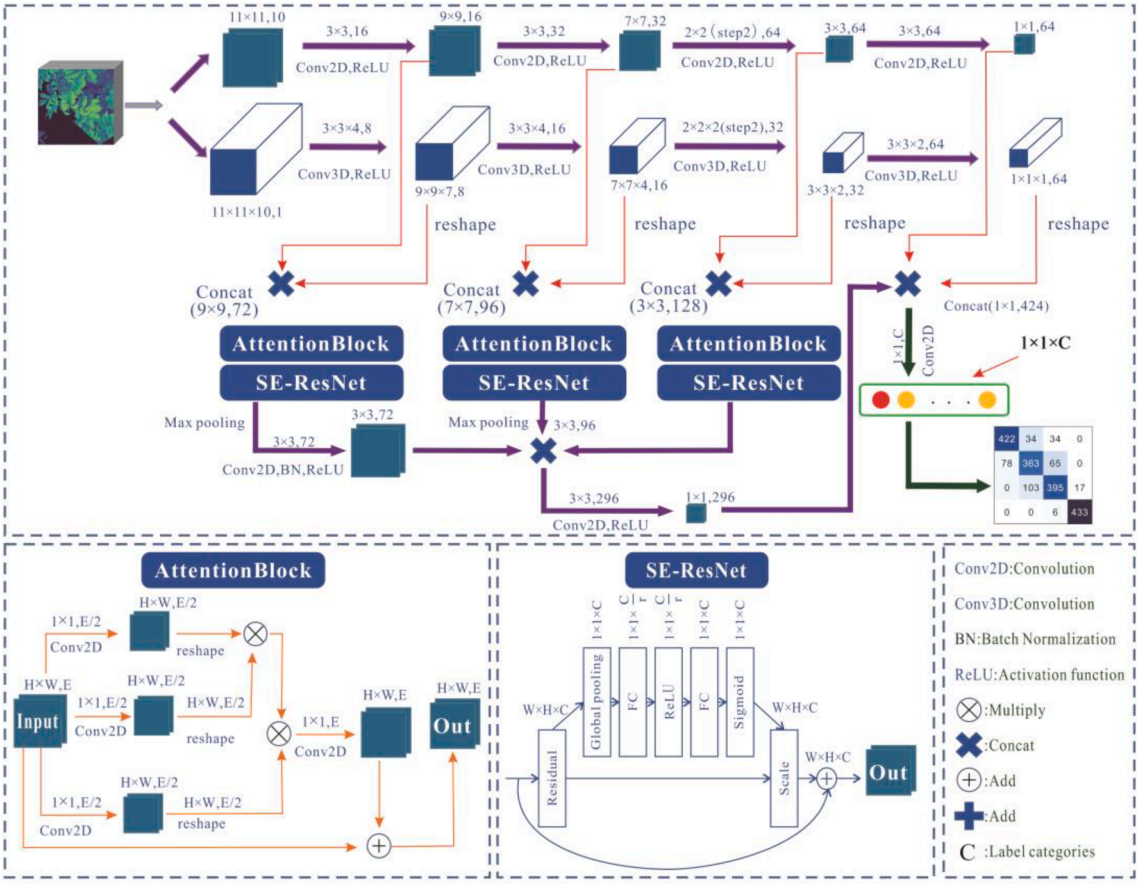


Fig. 4. The PLB-2D-3D-A structure. We first described the outline structure of Fig. 4 in Section 2.5.1. Secondly, we introduced how to combine 2D CNN and 3D CNN in Fig. 4 to make the feature map contain rich spatial-spectral correlation information in Section 2.5.2. Not only that, we presented the working mechanism of AttentionBlock and SE-ResNet in Fig. 4 in detail in Sections 2.5.3 and 2.5.4, respectively.

convolution kernel size of 1 × 1 and the number of convolution layers is $E/2$. Then, through two matrix multiplication operations and one regular softmax operation, the dimensions of the feature map obtained are $h \times w \times E/2$ (where h and w represent the width and height of X_{in}, respectively). The latitude of the feature map obtained using E Conv2D with a spatial kernel size of 1 × 1 to do the spatial convolution operation on the feature map is $h \times w \times E$. Finally, the obtained feature map is added with X_{in} to get the output O_{out}.

In contrast to convolution and pooling operations, AttentionBlock considers the relationship between distant pixels with the weights of all positions in the feature map. AttentionBlock focuses on the relevance between pixels in the whole feature map, whereas the convolution and pooling operations only concentrate on pixels in the spatial range of the convolution kernel size. The spatial size of the fused feature map is 9 × 9, 7 × 7, and 3 × 3. Since the spatial range is not very large, thus using AttentionBlock will only incur little computational complexity.

2.5.4. SE-ResNet

For CNN networks, the core computation is the convolution operator, which learns from the input data to the feature maps by employing a series of convolution kernels. In essence, convolution is the fusion of features over a local region, which includes spatially (H and W dimensions) as well as inter-channel (C dimension) fusion of features. For convolution operations, a large part of the work is to improve the perceptual field, i.e., to fuse more features over space, or to extract multi-scale spatial information, such as the multi-branch structure of Inception networks (Szegedy et al., 2016). For feature fusion in channel dimension, the convolution operation fuses all channels of the input feature map by default. The innovation of the SENet network (Li et al., 2021) is to focus on the relationship between channels, hoping that the model can automatically learn the importance of different channel features. For this purpose, SENet proposes the Squeeze-and-Excitation (SE) module, as shown in Fig. 4.

The SE module first performs the Squeeze operation on the feature map resulting from the convolution to acquire the global features at the channel level, followed by an excitation operation on the global features to learn the relationship between each channel, while obtaining the weights of the different channels, and finally multiplies the original feature map to obtain the final features. Essentially, the SE module is doing the attention or the rating operation on the channel dimension. This attention mechanism allows the model to pay more attention to the most informative channel features and suppress those unimportant channel features. Another point is that the SE module is generic, which means it can be integrated into existing network architectures.

The SE module mainly comprises two operations, Squeeze and Excitation (Hu et al., 2020), which can be adapted to any mapping. Take convolution as an example, the convolution kernel is $V = [v_1, v_2, \cdots, v_C]$, where v_C denotes the C_{th} convolution kernel. The output $U = [u_1, u_2, \cdots, u_C]$.

$$u_c = v_c * X = \sum_{s=1}^{C} v_c^s * x^s \tag{13}$$

Where * stands for the convolution operation, X represents the target image, and v_c^s refers to a 2D convolution kernel of an s channel, whose input is the spatial features on a channel, and it learns the feature spatial relations, but since the sum is done on the convolution results of each channel, the channel feature relations are mixed with the spatial relations learned by the convolution kernel. The SE module is designed to abstract away this mixing, so that the model can directly learn the channel feature relations.

Since convolution only operates in a local space, it is difficult for U to obtain enough information to extract the relationship between channels, which is more serious for the front layers of the network, because the perceptual field is relatively small. SENet proposes the Squeeze operation, which encodes the entire spatial feature on a channel as a global feature, using global average pooling to achieve.

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i,j), z \in R^C \tag{14}$$

H represents the height of the feature map. W stands for the width of the feature map.

After the Squeeze operation retrieves the global description of the features, we need another operation to capture the relationships between channels. This operation needs to satisfy-two criteria: firstly, it has to be flexible, and able to learn the nonlinear relationships between the individual channels; secondly, the learned relationships are not mutually exclusive, because here multi-channel features are allowed instead of the one-hot form. Based on this, the gating mechanism in sigmoid form is used here:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \mathrm{Re}LU(W_1 z)) \tag{15}$$

Where $W_1 \in R^{\frac{C}{r} \times C}$ and $W_2 \in R^{C \times \frac{C}{r}}$. To reduce the model complexity as well as to improve the generalization ability, a bottleneck structure containing two fully connected layers is adopted here, where the first FC layer plays the role of dimensionality reduction, and the dimensionality reduction factor of r is a hyperparameter, and then ReLU activation (Eckle and Schmidt-Hieber, 2019) is applied. The final FC layer restores the original dimensionality. Finally, the learned activation values (sigmoid activation (Han and Moraga, 1995), values 0 to 1) of each channel are multiplied by the original features on U:

$$\widetilde{x}c = Fscale(u_c, s_c) = s_c \bullet u_c \tag{16}$$

In fact, the whole operation can be seen as learning the weight coefficients of each channel, thus making the model more discriminative of the features of each channel, which should also be considered an attention mechanism. SE module applied in ResNet (Zhang et al., 2021), the model parameters and the amount of computation will increase, here take SE-ResNet-50 as an example, for the increase of model parameters is

$$\frac{2}{r} \sum_{s=1}^{S} N_s \bullet C_s^2 \tag{17}$$

where r denotes the number of descending coefficients, S indicates the number of stages, C_s refers to the number of channels in the s^{th} stage, and N_s stands for the amounts of duplicate blocks in the sth stage. When $r = 16$, SE-ResNet-50 only increases the number of parameters by about 10 %. However, the computational volume (GFLOPS (Vuduc et al., 2010)) is increased by less than 1 %.

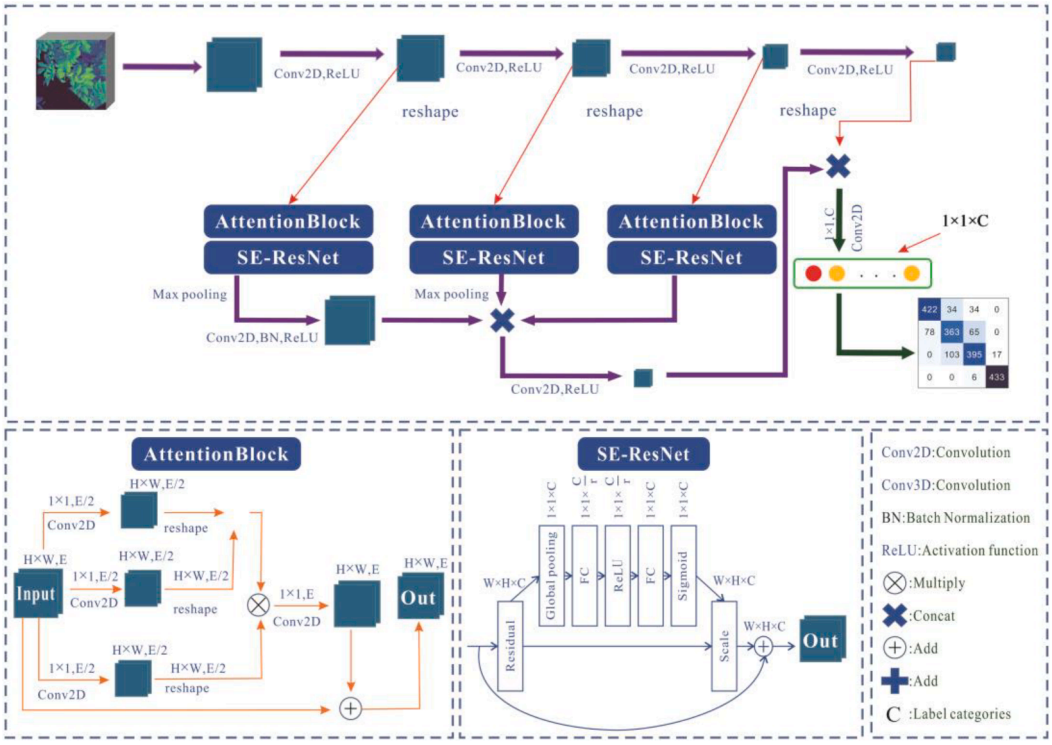
2.6. RF, 2D-CNN and 3D-CNN

Random Forest (RF) (Li et al., 2019) is a machine learning algorithm that combines hundreds of decision trees, where each tree depends on the values of independently sampled random vectors. Predictions are aggregated for classification or averaged for regression by majority voting on the predictions of the set. Since RF is a non-parametric method, it does not require values to follow a specific statistical distribution.

PLB-2D-3D-A is a classification algorithm based on 2D-CNN in cooperation with 3D-CNN, we split PLB-2D-3D-A into a 2D-CNN based classification algorithm and a 3D-CNN based classification algorithm, in the splitting process, we only change the feature extraction part, that is, we keep the whole network having an only 2D-CNN structure or 3D-CNN structure, the specific structure is shown in Fig. 5.

2.7. Evaluation metrics

Based on the confusion matrix, we obtained the following metrics: Accuracy, Precision, Recall and F1 Score (Jozdani and Chen, 2020). the



2D-CNN structure

Fig. 5. Customized 2D-CNN structure and 3D-CNN structure.

calculation formula is as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \tag{18}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{19}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{20}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{21}$$

Where TP, FP, TN and FN are the true positive, false positive, true negative and false negative samples, respectively.

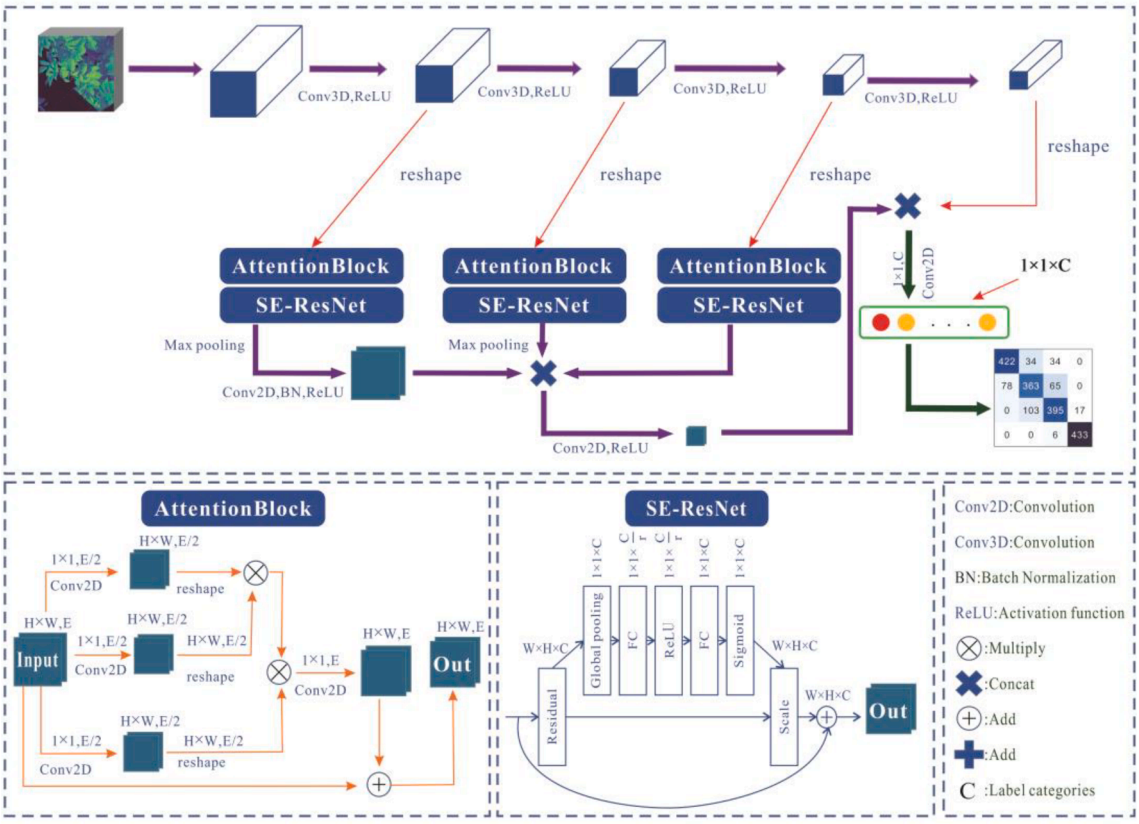
2.8. Experimental setup

We selected full-band samples and six specific bands from the band screening in Section 2.4 as input samples, respectively, to verify the classification effectiveness of model PLB-2D-3D-A for the early PLB. To investigate the detection performance of PLB-2D-3D-A in-depth, we compare the training results with the performance of three classification models (RF, 2D-CNN and 3D-CNN).

3. Results

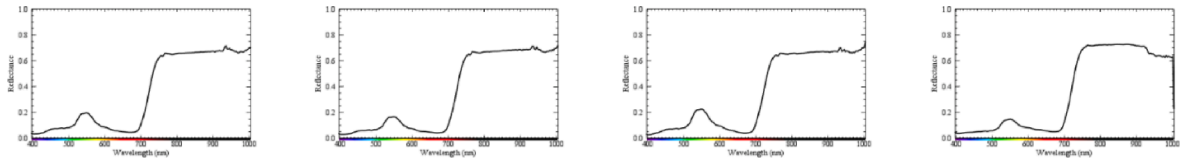
The spectral curves for 20 infected potato genotypes are shown in Fig. 6. According to the methodology (calculations 11 and 12) in Section 2.4, the band screening results are shown in Fig. 7 (a), where the peaks and troughs represent specific bands. Taking the intersection of all specific bands, we screened six important bands (492 nm, 519 nm, 560 nm, 592 nm, 717 nm and 765 nm). Moreover, we used the principal component analysis (PCA) (Tongcham et al., 2020) and the Pearson correlation coefficient (PCC) (Khosravi et al., 2019) to verify the feasibility of the approach used in this paper, and the results are shown in Fig. 7 (b). It can be observed that the wavelengths screened by the PCA and the PCC are basically around 492 nm, 519 nm, 560 nm, 592 nm, 717 nm and 765 nm (channel 34, channel 43, channel 57, channel 68, channel 110 and channel 126). From the experimental results, we found that the spectral first-order differentiation and spectral second-order are highly similar to PCA and PCC in terms of screening wavelengths. However, the methods we used are faster and more concise.Fig. 7a. Fig. 7b..

Table 1 shows the classification results of the four models under the full bands, and Fig. 8 shows the corresponding confusion matrix. It can be seen that PLB-2D-3D-A has a significant performance advantage, reaching an accuracy of 0.739. As expected with increased infection, the fourth class (July 18) has the best classification results, with precision, recall and F1 can reach 0.838, 0.892 and 0.864, respectively. Surprisingly, the first class (July 9) was second only to the fourth in terms of classification, with precision, recall and F_1 reaching 0.764, 0.830 and 0.796, respectively. The inoculation was on July 7, which means that on the third day of infection, the handheld hyperspectral camera was already able to capture the reflectance changes of the potato leaves. However, the classification performance of the second class (July 13) and the third class (July 15) is somewhat disappointing, especially for the second class, where precision, recall and F_1 only reach 0.632, 0.623 and 0.627, respectively, as seen from the confusion matrix in Fig. 8. Not only the PLB-2D-3D-A model, but also the RF, 2D-CNN and 3D-CNN have similar issues, with a large number of features entangled in the second and third classes. The RF achieves the lowest accuracy (0.683), probably because it fails to learn high-level complex image features in-

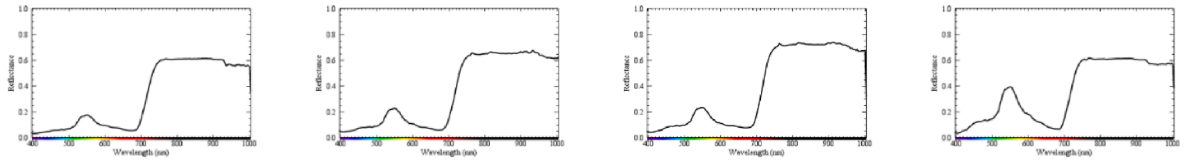


3D-CNN structure

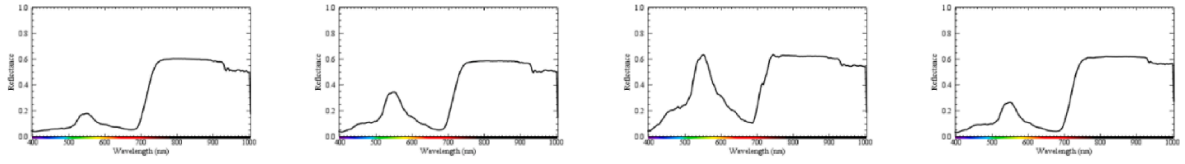
Fig. 5. (*continued*).



breeding line 9 breeding line 10 breeding line 11 Oleva



Bintje Danva breeding line 12 breeding line 13



breeding line 14 breeding line 15 Craigs Snow White breeding line 16

Fig. 6. Results of band screening.

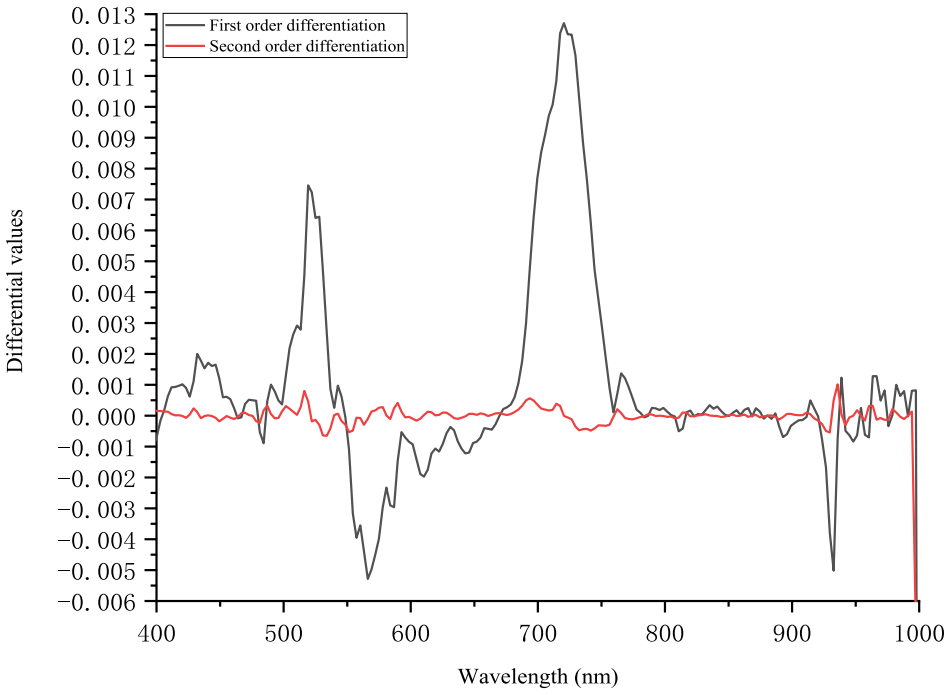
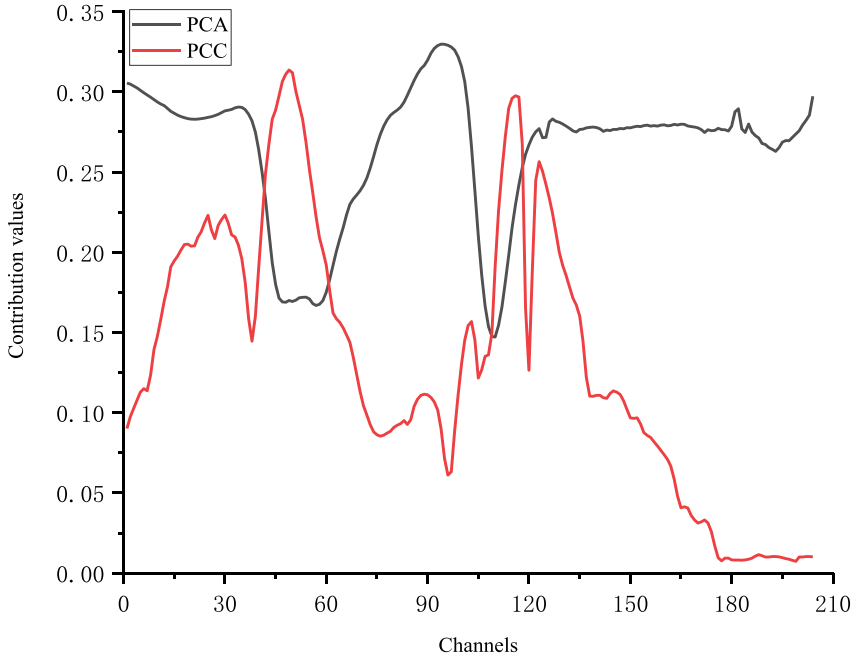


Fig. 7a. Band screening results of first and second order differentiation.



(b). Band screening results of the PCA and the PCC.

Fig. 7b. Band screening results of the PCA and the PCC.

Table 1
The classification results of the four models under the full bands.

Models		July9			July13			July15			July18		
	Accuracy	precision	recall	F1	precision	recall	F1	precision	recall	F1	precision	recall	F1
RF	0.683	0.696	0.825	0.755	0.576	0.535	0.555	0.680	0.571	0.621	0.778	0.874	0.823
2D-CNN	0.705	0.722	0.830	0.772	0.592	0.566	0.579	0.698	0.596	0.643	0.806	0.884	0.843
3D-CNN	0.717	0.730	0.820	0.772	0.604	0.584	0.594	0.710	0.616	0.660	0.822	0.890	0.855
Ours	0.739	0.764	0.830	0.796	0.632	0.623	0.627	0.722	0.638	0.677	0.838	0.892	0.864

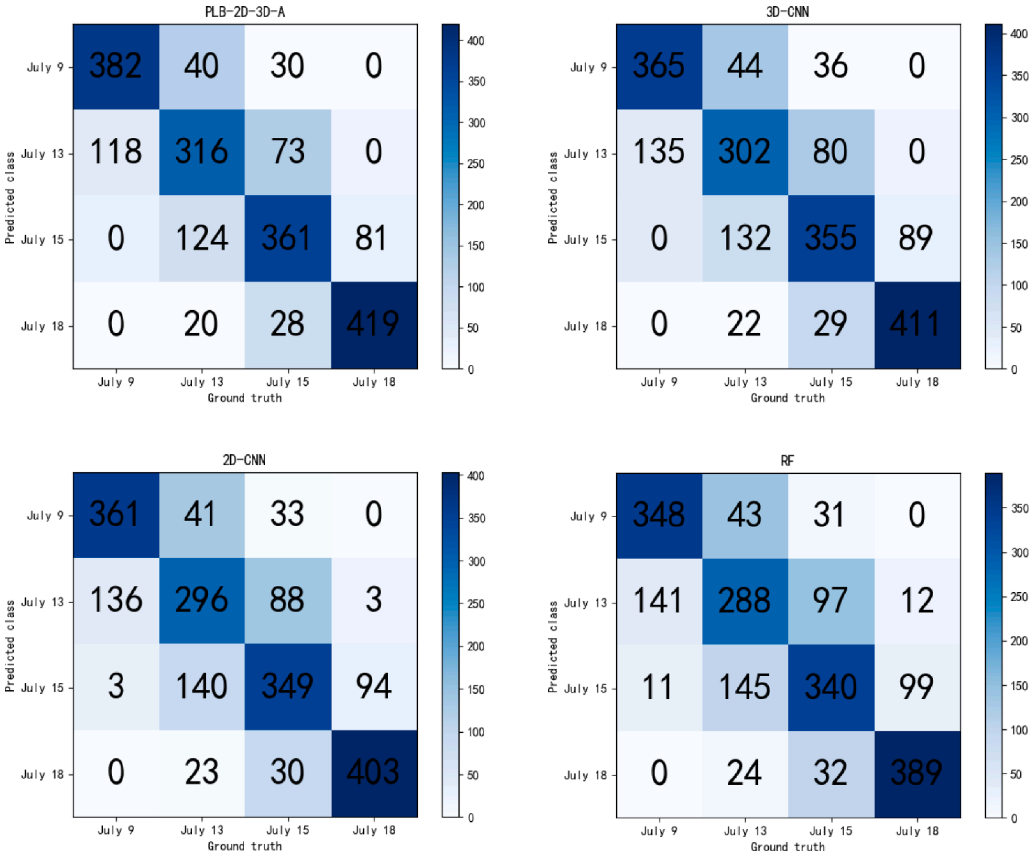


Fig. 8. Confusion matrix results of the four models at full bands.

depth. 3D-CNN has a slightly better performance than 2D-CNN, showing that the accurate acquisition of spatial information plays a key role in analyzing hyperspectral images.

Table 2 shows the classification results of the four models under six specific bands, and Fig. 9 shows the corresponding confusion matrices. When we select specific bands with high reflectance, the overall performance of the model is slightly improved, with RF, 2D-CNN, 3D-CNN and PLB-2D-3D-A Accuracy improved by 5.56 %, 6.38 %, 6.83 % and 6.9 %, respectively. It can be seen that the selection of specific bands is effective in improving classification performance.

To investigate the effect of attentional mechanisms on the proposed PLB-2D-3D-A, we designed ablation experiments to validate the AttentionBlock module and the SE-ResNet module. Specifically, we removed the AttentionBlock module and the SE-ResNet module in turn and tested them at full and specific bands, respectively, with the experimental setup remaining consistent with the above experiments. The detailed data of the results are shown in Table 3 and Table 4, and the confusion matrices in the test dataset are shown in Fig. 10 and Fig. 11.

It can be clearly observed that the attention mechanism has a positive effect on the classification performance. At the full band, removing the AttentionBlock module and the SE-ResNet module reduced the accuracy by 0.011 and 0.02, respectively. At the specific band, removing the AttentionBlock module and the SE-ResNet module reduced the accuracy by 0.011 and 0.021, respectively. It is worth noting that the AttentionBlock module has the most significant ability to distinguish the second class (July13). At specific bands, removing the AttentionBlock module reduced the precision, recall and F_1 of the model by 0.016. The SE-ResNet module showed the superior ability to distinguish the third category (July15). At specific bands, removing the SE-ResNet module reduced the precision, recall and F_1 of the model by 0.022, 0.037 and 0.031, respectively.

4. Discussion

This paper focuses on different stages detection of late blight caused by *P. infestans* in potato. It evaluates the combination of 2D-CNN and 3D-CNN with deep synergy. First, PLB-2D-3D-A uses a multilayer fusion strategy to design the network structure. In the horizontal structure, 2D convolution kernel and 3D convolution kernel are used to conduct shallow feature extraction for spectral features and spatial features respectively. In the vertical structure, the general convolution and pooling operations only consider pixels of kernel size and ignore the relationship between the whole feature image pixels, especially the relationship between two pixels that are far away from each other. This could lead to massive entanglement of similar features and seriously affect the classification ability of the model. Therefore, we introduce the

Table 2
The classification results of the four models under the specific bands.

Models		July9			July13			July15			July18		
	Accuracy	precision	recall	F1	precision	recall	F1	precision	recall	F1	precision	recall	F1
RF	0.721	0.734	0.889	0.804	0.674	0.589	0.629	0.704	0.584	0.638	0.772	0.937	0.847
2D-CNN	0.750	0.762	0.884	0.829	0.686	0.616	0.649	0.734	0.637	0.682	0.818	0.938	0.874
3D-CNN	0.766	0.774	0.882	0.824	0.704	0.648	0.675	0.748	0.658	0.700	0.836	0.929	0.880
Ours	0.790	0.796	0.888	0.839	0.720	0.679	0.699	0.764	0.700	0.731	0.878	0.922	0.899

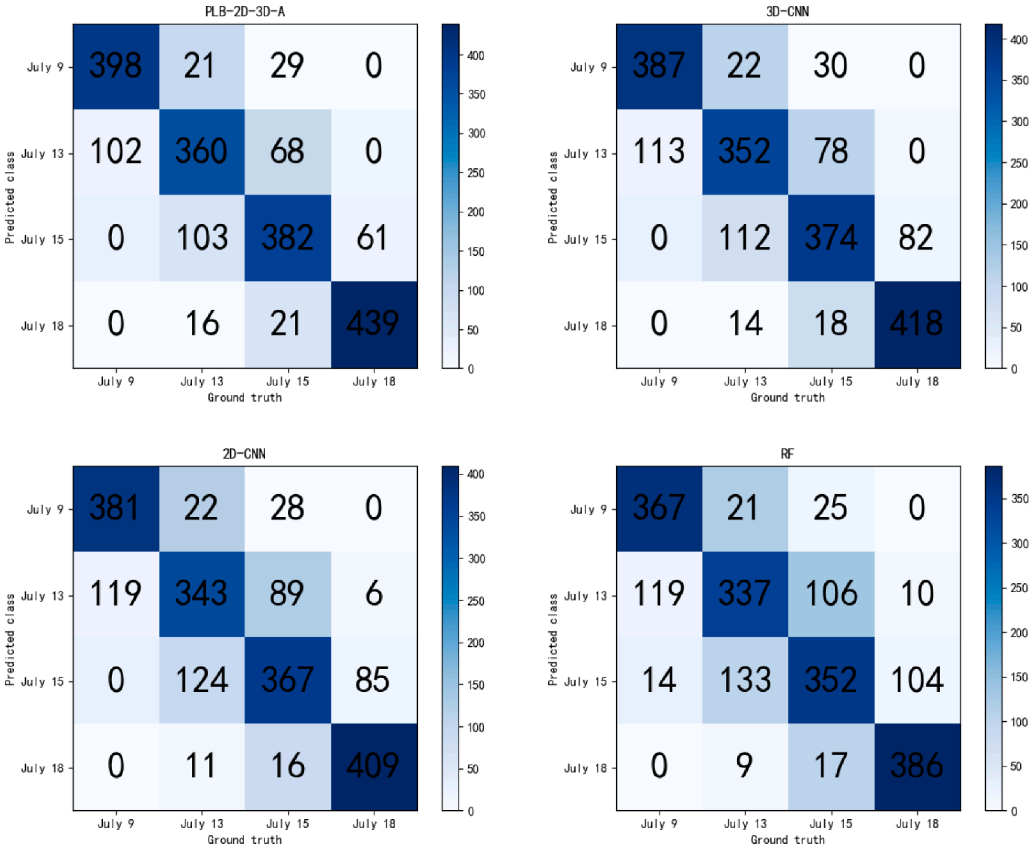


Fig. 9. Confusion matrix results of the four models at specific bands.

Table 3
Ablation experiments at full bands.

Models			July9			July13			July15			July18	
	Accuracy	precision	recall	F1	precision	recall	F1	precision	recall	F1	precision	recall	F1
Ours-A-S	0.700	0.708	0.829	0.764	0.580	0.541	0.560	0.702	0.597	0.645	0.808	0.900	0.852
Ours-S	0.719	0.744	0.842	0.790	0.612	0.583	0.597	0.704	0.600	0.648	0.814	0.913	0.861
Ours-A	0.728	0.752	0.836	0.792	0.622	0.599	0.610	0.710	0.616	0.660	0.826	0.908	0.865
Ours	0.739	0.764	0.830	0.796	0.632	0.623	0.627	0.722	0.638	0.677	0.838	0.892	0.864

Table 4
Ablation experiments at specific bands.

Models			July9			July13			July15			July18	
	Accuracy	precision	recall	F1	precision	recall	F1	precision	recall	F1	precision	recall	F1
Ours-A-S	0.741	0.742	0.877	0.804	0.680	0.605	0.640	0.720	0.632	0.673	0.820	0.921	0.868
Ours-S	0.769	0.780	0.894	0.833	0.700	0.643	0.670	0.742	0.663	0.700	0.852	0.926	0.887
Ours-A	0.779	0.786	0.889	0.834	0.704	0.663	0.683	0.760	0.677	0.716	0.866	0.929	0.895
Ours	0.790	0.796	0.888	0.839	0.720	0.679	0.699	0.764	0.700	0.731	0.878	0.922	0.899

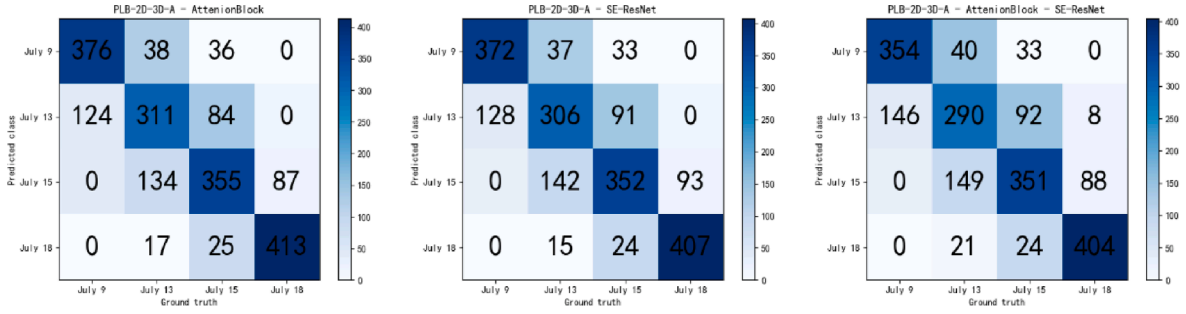


Fig. 10. Confusion matrix for ablation experiments at the full bands.

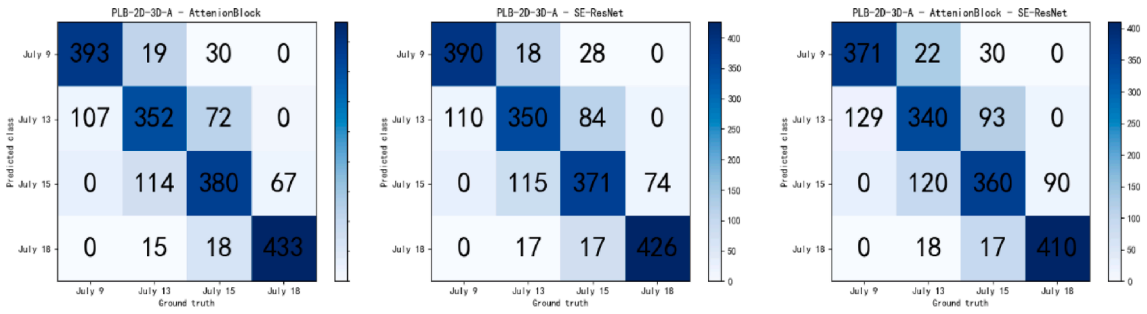


Fig. 11. Confusion matrix for ablation experiments at the specific bands.

AttentionBlock attention mechanism to provide a decoupling-like operation for similar features. However, applying the attention mechanism to the entire feature map results in a severe computational burden. To deal with this, we cropped the size of the original image and used small neighbourhood modules of different sizes. As shown in Table 3 and Table 4, by removing the AttentionBlock attention mechanism from the proposed model structure, the classification accuracy suffers both using all bands and specific bands. Specially for the second class (13 July), under the specific bands, the precision and recall were reduced by 2.22 % and 2.36 %, respectively. Due to the large number of sample parameters (over 50 million parameters to be processed for a single sample), we designed the horizontal and vertical structures with fewer convolution operations to ensure efficient network training, resulting in a shallow network depth and width. Generally, it is considered that increasing the depth and width of the network can effectively improve the model accuracy, such as Inception V3 and Resnet101 (Liao et al., 2020). To extract deeper spectral and spatial features, we use the widely accepted residual network as the underlying architecture and add the SE block for HSI classification after the residual transformation. The main role of the SE block is to recalibrate the feature mapping after the residual transformation. The SE block takes the feature map as its input and decomposes the spatial dimensional dependencies by global averaging pooling to learn a channel-oriented descriptor that it passes through a squeeze function. The goal of the descriptor is to emphasise useful channels by recalibrating the feature map, thus embedding the global distribution of feature maps for different channels, and is able to slightly improve the quality of the features generated by the residual block through explicitly modelling the relationship between the channels of its convolutional feature map. As shown in Table 3 and Table 4, when we remove the SE-ResNet module from the proposed model structure, the classification accuracy suffers both at full band and at specific bands. Especially for the third class (July 15), the precision and recall are reduced by 2.88 % and 2.36 % at the specific band, respectively. It can be seen that the AttentionBlock module and the SE-ResNet module can better fuse spatial and spectral features and have some non-negligible positive effects on the classification of infected samples.

After discussing the PLB-2D-3D-A network structure, we review the whole experimental part. First, to compare the classification effects at the full bands and the specific bands, we used spectral first-order differentiation and spectral second-order differentiation to screen out six specific bands (492 nm, 519 nm, 560 nm, 592 nm, 717 nm and 765 nm) through the intersections of the two differentiation results. We counted the mean values of the spectral curves of infected samples for four days, as shown in Fig. 12. Between 500 nm and 600 nm, the reflectance showed a clear regular variation as the infection level increased, consistent with the results of the band screening. When the band was between 700 nm and 800 nm, the spectral mean reflectance of the second class (July 13) and third class (July 15) showed confusion, especially the spectral curve of the second class (July 13). This is consistent with the results in Tables 1 and 2. At the full bands, the classification accuracy and recall of the second class (July 13) were 0.632 and 0.623, respectively. At the specific bands, the classification accuracy and recall

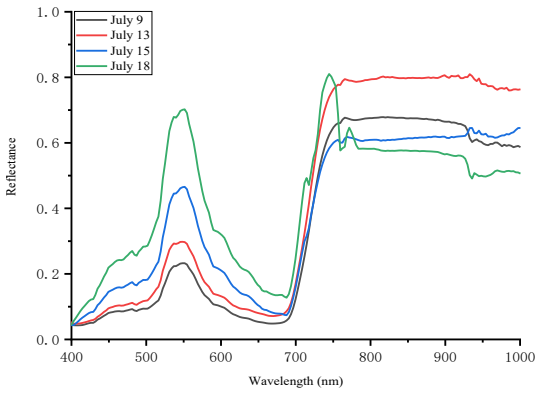


Fig. 12. Spectral curves of infected samples for four days.

of the second class (July 13) were 0.720 and 0.679, respectively. when we removed the AttentionBlock module and the SE-ResNet module entirely, as shown in Table 3 and Table 4, the classification accuracy and recall of the second class (July 13) decreased by 8.23 % and 13.16 %, respectively. At the specific bands, the classification accuracy and recall of the second class (July 13) decreased by 5.56 % and 10.9 %, respectively. At the full band, the classification accuracy and recall of the third class (July 15) decreased by 2.86 % and 6.43 %, respectively. At the specific band, the classification accuracy and recall of the second class (13 July) sample decreased by 5.76 % and 9.71 %, respectively. The positive impact of the AttentionBlock module and the SE-ResNet module on HSI classification was verified from a quantitative perspective. An interesting observation is that when removing the AttentionBlock module at full bands, the recall improves by 0.016 and when removing the SE-ResNet module, the recall improves by 0.021. When removing both the AttentionBlock and SE-ResNet modules, the recall improves by 0.008. It should be stated that the goal of this paper is to address the early detection of late blight, and the fourth class of samples already has some lesions that could be captured by the experienced human eye, the fourth class is not considered an early detection of late blight. However, we found that the attention mechanism was not effective in classifying the class with the strongest variation in reflectance, probably because, the features of the fourth class could be extracted well by the traditional convolution mechanism, and embedding the attention mechanism would conversely generate redundant invalid features, thus reducing the generalization ability of the model. When the band is between 800 nm and 1000 nm, the reflectance again shows some degree of regularity. When the band is between 910 nm and 940 nm, the first class (July 9) and the third class (July 15) show a crossover point, indicating that these two classes are prone to feature entanglement. Some studies (Rodriguez et al., 2021) have shown that when using individual cultivars, there will be some improvement in classification performance in a specific band. When modelling for one cultivar, we need to avoid using feature wavelengths with potentially substantial feature entanglement. It is

worth emphasising that by using 20 cultivars and two band screening methods, the final results are intersected and many species-specific bands are removed, for example in breeding line 1, 463 nm is a specific band and when intersected, this specific band is removed. This will result in a slight decrease in the accuracy of the model. In a realistic scenario, it is not practical to model each cultivar individually and the future trend is to integrate more cultivars for modelling. In general, PLB-2D-3D-A performed the worst classification for the second class and the first class was second only to the fourth class. This indicates that it is possible to detect the early stages of late blight by capturing variations in spectral information on the third day of infection with *P infestans* in a field environment. It is unfair to direct compare the results in this paper with other study since different datasets were used, but it still shows a promising accuracy (0.796) and recall (0.888) values for in-field recognition (July 9) of PLB with the six selected important bands (492 nm, 519 nm, 560 nm, 592 nm, 717 nm and 765 nm) based proposed PLB-2D-3D-A model.

The red edge is the wavelength between 680 nm and 750 nm where the slope of the plant reflectance curve is maximum, i.e. the variation in reflectance from the low value region of red light (caused by chlorophyll absorption) to the high value region (leaf and canopy scattering effects). We calculated the red-edge parameter (Ruiliang et al., 2003) by calculating the first-order differentiation of the spectral reflectance between 680 nm and 750 nm, and the results are shown in Fig. 13. We found that the peak first-order differential between the 680 nm and 750 nm bands slowly shifted towards the longer wavelengths with the severity of the infection level, and a red-edge displacement phenomenon occurred. Existing studies have found that changes in chlorophyll and nitrogen with the health of the vegetation can cause red-edge displacement (Molina-Bolivar et al., 2019) phenomena in the spectrum. The occurrence of red-edge shifts in infected potato leaves also validates from a spectrochemical point of view that it is possible to achieve late blight detection.

The classification performance of the PLB-2D-3D-A model for the asymptomatic biological phase of potato late blight shows some encouraging results, but is not convincing enough. In future work, further improvements in model accuracy, particularly for class 2 and class 3, is an urgent matter that needs to be addressed. Not only that, the dataset in this paper does not involve healthy samples, and distinguishing healthy samples from early infected samples is a necessary task to explore. Furthermore, canopy level detection over large areas can significantly improve operational efficiency in real-world scenarios, whereas this paper focuses on canopy level detection in the canopy of individual plants. Therefore, it is also meaningful to establish an accurate geometric-optical model between canopy and leaf spectra. Finally, the emergence of the red-edge displacement phenomenon has brought us inspiration. With a multispectral camera mounted to a UAV, we previous picked up this change in the red-edge range at the canopy level in a later stage if infection in the EnBlightMe project (Chawade et al., 2019). Currently our work is mainly focused on the detection of late blight, and the understanding between the infection development could be improved by accompanying hyperspectral measurements with invasive methods such as pathogen detection by molecular techniques or microscope monitoring of the pathogen. It is worth exploring the possibility of adding factors to the deep learning inputs that could significantly influence the variations in spectral reflectance, such as variations in chlorophyll, nitrogen content and plant development, to further improve the model by incorporating quantitative conditioning mechanisms.

5. Conclusions

This paper proposes a deep learning architecture combining 2D-CNN and 3D-CNN, called PLB-2D-3D-A, for classification of the asymptomatic biotrophic phase of PLB disease. We collected the original dataset of 20 potato genotypes with 15,360 (64x64x204) images. The proposed model achieved an accuracy of 0.739 at the full bands in the test dataset. The accuracy and recall on day 12 after inoculation were the highest, reaching 0.838 and 0.892, respectively. The accuracy and recall on day 3 were second only to the day 12, reaching 0.764 and 0.830, respectively. The accuracy and recall on day 7 were the lowest, reaching 0.632 and 0.623, respectively. The proposed model achieves an accuracy of 0.790 at the specific bands (492 nm, 519 nm, 560 nm, 592 nm, 717 nm and 765 nm). The accuracy and recall on day 12 are the highest, reaching 0.878 and 0.922, respectively. The accuracy and recall on day 3 are second only to the day 12, reaching 0.796 and 0.888, respectively. The accuracy and recall on day 7 are the lowest, reaching 0.720 and 0.679, respectively. The results show that both in full bands and in specific bands, the proposed model outperforms the accuracy of the traditional machine learning method RF and the deep learning-based methods 2D-CNN and 3D-CNN. The study shows that our deep learning-based model achieves a promising result for in-field classification of the asymptomatic biotrophic phase of PLB based on proximal hyperspectral images.

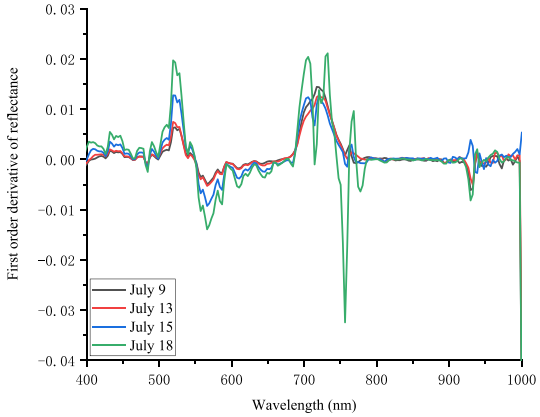


Fig. 13. Mapping relationship between wavelength and reflectance first order differentiation.

CRediT authorship contribution statement

Chao Qi: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Murilo Sandroni:** Writing – review & editing. **Jesper Cairo Westergaard:** Writing – review & editing. **Ea Høegh Riis Sundmark:** Resources, Writing – review & editing. **Merethe Bagge:** Resources, Writing – review & editing. **Erik Alexandersson:** Supervision, Project administration, Funding acquisition, Writing – review & editing. **Junfeng Gao:** Supervision, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We would like to thank Mathias Timmerman (Danespo) for his assistance. This research was founded by the Nordic Council of Ministers, Copenhagen, Denmark (PPP #6P2 and #6P3) and NordForsk, Norway (#84597). This work was partially supported by the European Union's Horizon 2020 (H2020) Marie Skłodowska-Curie Actions (grant

agreement number 766048) and Lincoln Agri-Robotics as part of the Expanding Excellence in England (E3) Programme.

References

Bagheri, N., 2020. Application of aerial remote sensing technology for detection of fire blight infected pear trees. Comput. Electronics Agriculture. 168, 105147 https://doi.org/10.1016/j.compag.2019.105147.

Behmann, J., Acebron, K., Emin, D., Bennertz, S., Matsubara, S., Thomas, S., Bohnenkamp, D., Kuska, M.T., Jussila, J., Salo, H., Mahlein, A.K., Rascher, U., 2018. Specim IQ: Evaluation of a New, Miniaturized Handheld Hyperspectral Camera and Its Application for Plant Phenotyping and Disease Detection. Sensors 18, 441. https://doi.org/10.3390/s18020441.

Chawade, A., van Ham, J., Blomquist, H., Bagge, O., Alexandersson, E., Ortiz, R.J.A., 2019. High-throughput field-phenotyping tools for plant breeding and precision agriculture. Agronomy 9, 258. https://doi.org/10.3390/agronomy9050258.

Chen, X.K., Wang, W.B., Cai, P.P., Wang, Z.W., Li, T.T., Du, Y., 2021. The role of the MAP kinase-kinase protein STMKK1 in potato immunity to different pathogens. Hortic. Res. 8, 9. https://doi.org/10.1038/s41438-021-00556-5.

Duarte-Carvajalino, J.M., Alzate, D.F., Ramirez, A.A., Santa-Sepulveda, J.D., Fajardo-Rojas, A.E., Soto-Suarez, M., 2018. Evaluating Late Blight Severity in Potato Crops Using Unmanned Aerial Vehicles and Machine Learning Algorithms. Remote Sens. (Basel). 10, 17. https://doi.org/10.3390/rs10101513.

Eckle, K., Schmidt-Hieber, J., 2019. A comparison of deep networks with ReLU activation function and linear spline-type methods. Neural Netw. 110, 232–242. https://doi.org/10.1016/j.neunet.2018.11.005.

Gao, J., Westergaard, J.C., Sundmark, E.H.R., Bagge, M., Liljeroth, E., Alexandersson, E., 2021. Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning. Knowl.-Based Syst. 214, 106723 https://doi.org/10.1016/j.knosys.2020.106723.

Gold, K.M., Townsend, P.A., Herrmann, I., Gevens, A.J., 2020. Investigating potato late blight physiological differences across potato cultivars with spectroscopy and machine learning. Plant Sci. 295, 12. https://doi.org/10.1016/j.plantsci.2019.110316.

Han, Z.Z., Deng, L.M., 2018. Application driven key wavelengths mining method for aflatoxin detection using hyperspectral data. Comput. Electron. Agric. 153, 248–255. https://doi.org/10.1016/j.compag.2018.08.018.

Han, J., Moraga, C., 1995. The influence of the sigmoid function parameters on the speed of backpropagation learning. Int. Workshop on artificial neural networks 930, 195–201. https://doi.org/10.1007/3-540-59497-3_175.

Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E., 2020. Squeeze-and-Excitation Networks. IEEE Trans. Pattern Anal. Mach. Intell. 42, 2011–2023. https://doi.org/10.1109/TPAMI.2019.2913372.

Ioffe, S. and Szegedy, C. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Batch normalization: Accelerating deep network training by reducing internal covariate shift, International conference on machine learning, 448-456: PMLR.

Jozdani, S., Chen, D.M., 2020. On the versatility of popular and recently proposed supervised evaluation metrics for segmentation quality of remotely sensed images: An experimental case study of building extraction. ISPRS J. Photogramm. Remote Sens. 160, 275–290. https://doi.org/10.1016/j.isprsjprs.2020.01.002.

Khosravi, K., Daggupati, P., Alami, M.T., Awadh, S.M., Ghareb, M.I., Panahi, M., Pham, B.T., Rezaie, F., Qi, C.C., Yaseen, Z.M., 2019. Meteorological data mining and hybrid data-intelligence models for reference evaporation simulation: A case study in Iraq. Comput. Electron. Agric. 167, 105041 https://doi.org/10.1016/j.compag.2019.105041.

Kingma, D. P. and Ba, J. J. 2014. Adam: A Method for Stochastic Optimization. arXiv [Preprint]. arXiv:1412.6980.

Krizhevsky, A., Sutskever, I., Hinton, G., 2012. Imagenet classification with deep convolutional neural networks. In: Adv. Neural Information Processing Systems, p. 25.

Li, Z., Jiang, K., Qin, S., Zhong, Y., Elofsson, A., 2021. GCSENet: A GCN, CNN and SENet ensemble model for microRNA-disease association prediction. PLoS Comput. Biol. 17, 6. https://doi.org/10.1371/journal.pcbi.1009048.

Li, F.S., Lehtomaki, M., Elberink, S.O., Vosselman, G., Kukko, A., Puttonen, E., Chen, Y., Hyyppa, J., 2019. Semantic segmentation of road furniture in mobile laser scanning data. ISPRS J. Photogramm. Remote Sens. 154, 98–113. https://doi.org/10.1016/j.isprsjprs.2019.06.001.

Liao, L., Zhao, Y., Wei, S., Wei, Y., Wang, J., 2020. Parameter Distribution Balanced CNNs. IEEE Trans. Neural Networks Learn. Syst. 31, 4600–4609. https://doi.org/10.1109/TNNLS.2019.2956390.

Ma, T., Tsuchikawa, S., Inagaki, T., 2020. Rapid and non-destructive seed viability prediction using near-infrared hyperspectral imaging coupled with a deep learning approach. Comput. Electron. Agric. 177 https://doi.org/10.1016/j.compag.2020.105683.

Meng, S.Y., Wang, X.Y., Hu, X., Luo, C., Zhong, Y.F., 2021. Deep learning-based crop mapping in the cloudy season using one-shot hyperspectral satellite imagery. Comput. Electron. Agric. 186, 105683 https://doi.org/10.1016/j.compag.2021.106188.

Molina-Bolivar, J.A., Galisteo-Gonzalez, F., Ruiz, C.C., Medina-O'Donnell, M., Martinez, A., Parra, A., 2019. Maslinic acid conjugate with 7-amino-4-methylcoumarin as probe to monitor the temperature dependent conformational changes of human serum albumin by FRET. Spectrochim. Acta Part a-Molecul. Biomolecul. Spectrosc. 214, 161–169. https://doi.org/10.1016/j.saa.2019.02.014.

Paoletti, M.E., Haut, J.M., Plaza, J., Plaza, A., 2018. A new deep convolutional neural network for fast hyperspectral image classification. ISPRS J. Photogramm. Remote Sens. 145, 120–147. https://doi.org/10.1016/j.isprsjprs.2017.11.021.

Rodriguez, J., Lizarazo, I., Prieto, F., Angulo-Morales, V., 2021. Assessment of potato late blight from UAV-based multispectral imagery. Comput. Electron. Agric. 184, 16. https://doi.org/10.1016/j.compag.2021.106061.

Ruiliang, P., Peng, G., Biging, G.S., Larrieu, M.R., 2003. Extraction of red edge optical parameters from Hyperion data for estimation of forest leaf area index. IEEE Trans. Geosci. Remote Sens. 41, 916–921. https://doi.org/10.1109/TGRS.2003.813555.

Savitzky, A. and Golay, M. J. J. A. c. 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry. 36, 1627-1639. doi: https://pubs.acs.org/doi/pdf/10.1021/ac60214a047.

Shi, Y., Han, L., Kleerekoper, A., Chang, S., Hu, T.J., 2021. A Novel CropdocNet for Automated Potato Late Blight Disease Detection from the Unmanned Aerial Vehicle-based Hyperspectral Imagery. arXiv[Preprint]. arXiv:2107.13277.

Su, J.Y., Yi, D.W., Su, B.F., Mi, Z.W., Liu, C.J., Hu, X.P., Xu, X.M., Guo, L., Chen, W.H., 2021. Aerial Visual Perception in Smart Farming: Field Study of Wheat Yellow Rust Monitoring. IEEE Trans. Ind. Inf. 17, 2242–2249. https://doi.org/10.1109/tii.2020.2979237.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In: ProceedIngs of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826.

Tongcham, P., Supa, P., Pornwongthong, P. and Prasitmeeboon, P. 2020. Mushroom spawn quality classification with machine learning. Computers and Electronics in Agriculture. 179. doi: 10, 105865. doi: 1016/j.compag.2020.105865.

Vuduc, R., Chandramowlishwaran, A., Choi, J., Guney, M., Shringarpure, A., 2010. On the limits of GPU acceleration. In: Proceedings of the 2nd USENIX Conference on Hot Topics in Parallelism, p. 13.

Wang, X.D., Zheng, K.K., Cheng, W.Y., Li, J., Liang, X.X., Shen, J., Dou, D.L., Yin, M.Z., Yan, S., 2021. Field application of star polymer-delivered chitosan to amplify plant defense against potato late blight. Chem. Eng. J. 417, 14. https://doi.org/10.1016/j.cej.2021.129327.

Yang, Y., Anderson, M.C., Gao, F., Wood, J.D., Gu, L.H., Hain, C., 2021. Studying drought-induced forest mortality using high spatiotemporal resolution evapotranspiration data from thermal satellite imaging. Remote Sens. Environ. 265, 14. https://doi.org/10.1016/j.rse.2021.112640.

Yin, S., Wang, Y., Yang, Y.H.J.P.R., 2020. A novel image-dehazing network with a parallel attention block. Pattern Recogn. 102, 107255 https://doi.org/10.1016/j.patcog.2020.107255.

Zhang, Y., Gao, J., Cen, H., Lu, L.Y., Yu, X.Y., He, Y., Pieters, J.G., 2019. Automated spectral feature extraction from hyperspectral images to differentiate weedy rice and barnyard grass from a rice crop. Comput. Electron. Agric. 159, 42–49. https://doi.org/10.1016/j.compag.2019.02.018.

Zhang, H., Jiang, L., Li, C., 2021. CS-ResNet: Cost-sensitive residual convolutional neural network for PCB cosmetic defect detection. Expert Syst. Appl. 185, 115673 https://doi.org/10.1016/j.eswa.2021.115673.

Zheng, K.K., Lu, J., Li, J., Yu, Y., Zhang, J., He, Z.W., Omayma, M., Wu, J., Xie, X.J., Li, X.B., Xu, G.Y., Dou, D.L., Wang, X.D., 2021. Efficiency of chitosan application against Phytophthora infestans and the activation of defence mechanisms in potato. Int. J. Biol. Macromol. 182, 1670–1680. https://doi.org/10.1016/j.ijbiomac.2021.05.097.